References

Abadie A, Chingos MM, West MR (2013). Endogenous Stratification in Randomized Experiments. NBER working paper no. 19742.

Abadie A, Imbens GW (2006). Large sample properties of matched estimates for average treatment effects. *Econometrica* 74:235–267.

Angrist JD, Imbens GW, Rubin DB (1996). Identification of causal effects using instrumental variables. *Journal of the American Statistical Association* 91:444–455.

Bang H, Robins JM (2005). Doubly robust estimation in missing data and causal inference models. *Biometrics* 61: 962-972 (errata appeared in *Biometrics* 2008;64:650).

Balke A, Pearl J (1994). Probabilistic evaluation of counterfactual queries. In *Proceedings of the Twelfth National Conference on Artificial Intelligence*, Volume I, pp. 230–237.

Berkson J (1946). Limitations of the application of fourfold table analysis to hospital data. *Biometrics* 2: 47–53.

Berkson J (1955). The statistical study of association between smoking and lung cancer. *Proceedings of the Staff Meetings of the Mayo Clinic* 30: 319–348.

Blot WJ, Day NE (1979). Synergism and interaction: are they equivalent? [letter] *American Journal of Epidemiology* 110:99–100.

Casella G, Berger RL (2002). *Statistical Inference*, 2nd ed. Pacific Grove, CA: Duxbury Press.

Cornfield J, Haenszel W, Hammond EC, Lilienfeld AM, Shimkin MB, Wynder EL (1959). Smoking and lung cancer: Recent evidence and a discussion of some questions. *Journal of the National Cancer Institute* 22:173–203. Reprinted in *International Journal of Epidemiology* 2009; 38(5): 1175-1191.

Cox DR (1958). *Planning of Experiments*. New York, NY: John Wiley and Sons.

Dawid AP (1979). Conditional independence in statistical theory (with discussion). *Journal of the Royal Statistical Society B* 41:1–31.

Dawid AP (2000). Causal inference without counterfactuals (with discussion). *Journal of American Statistical Association* 95: 407–424.

Dawid AP (2002). Influence diagrams for causal modelling and inference. *International Statistical Review* 70: 161–189.

de Finetti (1972). *Probability, Induction, and Statistics*. John Wiley & Sons.

Dosemeci M, Wacholder S, Lubin JH (1990). Does nondifferential misclassification of exposure always bias a true effect toward the null value? *American Journal of Epidemiology* 132:746–748.

Flanders WD (2006). On the relation of sufficient component cause models with potential (counterfactual) models. *European Journal of Epidemiology* 21: 847–853.

Fleming TR, Harrington DP (2005). *Counting Processes and Survival Analysis*, 2nd ed. New York: Wiley.

Gelman A, Carlin JB, Stern HS, Rubin DB (2003). *Bayesian Data Analysis*, 2nd ed. Boca Raton, FL: Chapman & Hall/CRC.

Glymour MM, Tchetgen Tchetgen EJ, Robins JM (2012). Credible Mendelian randomization studies: approaches for evaluating the instrumental variable assumptions. *American Journal of Epidemiology* 175: 332-339.

Greenland S (1987). Interpretation and choice of effect measures in epidemiologic analyses. *American Journal of Epidemiology* 125:761–768.

Greenland S (1996a). Basic methods for sensitivity analysis of bias. *International Journal of Epidemiology* 25:1107–1116.

Greenland S (1996b). Absence of confounding does not correspond to collapsibility of the rate ratio or rate difference. *Epidemiology* 7:498–501.

Greenland S (2003). Quantifying biases in causal models: classical confounding versus collider-stratification bias. *Epidemiology* 14:300–306.

Greenland S (2008). Introduction to regression modeling. In: Rothman KJ, Greenland S, Lash TL, eds. *Modern Epidemiology*, 3rd edition. Philadelphia, PA: Lippincott Williams & Wilkins, pp. 418–455.

Greenland S (2009a). Bayesian perspectives for epidemiologic research. III. Bias analysis via missing-data methods. *International Journal of Epidemiology* 38:1662–1673.

Greenland S (2009b). Relaxation penalties and priors for plausible modeling of nonidentified bias sources. *Statistical Science* 24:195–210.

Greenland S, Brumback B (2002). An overview of relations among causal modelling methods. *International Journal of Epidemiology* 31:1030–1037.

Greenland S, Lash TL (2008). Bias analysis. In: Rothman KJ, Greenland S, Lash TL, eds. *Modern Epidemiology*, 3rd edition. Philadelphia, PA: Lippincott Williams & Wilkins, pp. 345–380.

Greenland S, Lash TL, Rothman KJ (2008). Concepts of interaction. In: Rothman KJ, Greenland S, Lash TL, eds. *Modern Epidemiology*, 3rd edition. Philadelphia, PA: Lippincott Williams & Wilkins, pp. 71–83.

Greenland S, Robins JM (1986). Identifiability, exchangeability, and epidemiological confounding. *International Journal of Epidemiology* 15:413–19.

Greenland S, Robins JM (1988). Conceptual problems in the definition and interpretation of attributable fractions. *American Journal of Epidemiology* 128:1185–1197.

Greenland S, Robins JM (2009). Identifiability, exchangeability, and confounding revisited. *Epidemiologic Perspectives & Innovations* 6:4.

Greenland S, Pearl J, Robins JM (1999). Causal diagrams for epidemiologic research. *Epidemiology* 10:37–48.

Greenland S, Robins JM, Pearl J (1999). Confounding and collapsibility in causal inference. *Statistical Science* 14:29–46.

Greenland S, Rothman KJ (2008). Introduction to stratified analysis. In: Rothman KJ, Greenland S, Lash TL, eds. *Modern Epidemiology*, 3rd edition. Philadelphia, PA: Lippincott Williams & Wilkins, pp. 258–282.

Grieve AP (2003). The number needed to treat: a useful clinical measure or a case of the Emperor's new clothes? *Pharmaceutical Statistics* 2:87–102.

Halloran ME, Struchiner CJ (1995). Causal inference in infectious diseases. *Epidemiology* 6: 142–151.

Hastie TJ, Tibshirani RJ (1990). *Generalized Additive Models*. London: Chapman & Hall.

Hansen BB (2008). The prognostic analogue of the propensity score. *Biometrika* 95(2): 481–488.

Hernán MA, Clayton D, Keiding N (2011). The Simpson's paradox unraveled. *International Journal of Epidemiology* 40:780-785.

Hernán MA, Hernández-Díaz S, Werler MM, Mitchell AA (2002). Causal knowledge as a prerequisite for confounding evaluation: an application to birth defects epidemiology. *American Journal of Epidemiology* 155:176–184.

Hernán MA, Hernández-Díaz S, Robins JM (2004). A structural approach to selection bias. *Epidemiology* 15:615–625.

Hernán MA, Hernández-Díaz S (2012). Beyond the intention to treat in comparative effectiveness research. *Clinical Trials*; 9(1):48-55.

Hernán MA, Robins JM (2006). Estimating causal effects from epidemiological data. *Journal of Epidemiology and Community Health* 60: 578–586.

Hernán MA, Robins JM (2006b). Instruments for causal inference: An epidemiologist's dream? *Epidemiology* 2006; 17(4): 360–372.

Hernán MA, VanderWeele TJ (2011). Compound treatments and transportability of causal inference. *Epidemiology* 22:368–377.

Holland PW (1986). Statistics and causal inference (with discussion). *Journal of the American Statistical Association* 81:945–961.

Horvitz DG, Thompson DJ (1952). A generalization of sampling without replacement from a finite universe. *Journal of the American Statistical Association* 47:663–685.

Hosmer DW, Lemeshow S, May S (2008). *Applied Survival Analysis: Regression Modelling of Time to Event Data*. Hoboken, NJ: Wiley.

Hudgens MG, Halloran ME (2009). Towards causal inference with interference. *Journal of the American Statistical Association* 103:832–842.

Hume D (1748). *An Enquiry Concerning Human Understanding*. Reprinted and edited 1993, Indianapolis/Cambridge: Hacket.

Imbens GW, Angrist JD (1994). Identification and estimation of local average treatment effects. *Econometrica* 62:467–475.

Kalbfleisch and Prentice (2002). *The Statistical Analysis of Failure Time Data*. Hoboken, NJ: Wiley.

Kosinski S, Stillwell D, Graepel T (2013). Private traits and attributes are predictable from digital records of human behavior. *Proceedings of the National Academy of Sciences* 110(15): 5802–5805.

Kurth T, Walker AM, Glynn RJ, Chan KA, Gaziano JM, Berger K, Robins JM (2006). Results of multivariable logistic regression, propensity matching, propensity adjustment, and propensity-based weighting under conditions of nonuniform effect. *American Journal of Epidemiology* 163(3): 262–270.

Laupacis A, Sackett DL, Roberts RS (1988). An assessment of clinically useful measures of the consequences of treatment. *New England Journal of Medicine* 318:1728–1733.

Lauritzen SL, Dawid AP, Larsen BN, Leimer H-G (1990). Independence properties of directed Markov fields. *Networks* 20:491–505.

Lash TL, Fox MP, Fink AK (2009). *Applying Quantitative Bias Analysis to Epidemiologic Data*. New York: Springer.

Lewis D (1973). *Counterfactuals*. Oxford: Blackwell.

Liang K-Y, Zeger SL (1986). Longitudinal data analysis using generalized linear models. *Biometrika* 73(1):13-22.

Mackie JL (1965). Causes and conditions. *American Philosohpical Quartely* 2:245-264.

Manski CF (1990). Nonparametric bounds on treatment effects. *American Economic Review* 80(2):319-323.

McCullagh P, Nelder JA (1989). *Generalized Linear Models*, 2nd ed. London: Chapman & Hall.

McCulloch CE, Searle SE, Neuhaus JM (2008). *Generalized, Linear, and Mixed Models*, 2nd ed. New York: Wiley.

Miettinen OS (1973). Standardization of risk ratios. *American Journal of Epidemiology* 96:383–388.

Miettinen OS (1982). Causal and preventive interdependence: elementary principles. *Scandinavian Journal of Work, Environment & Health* 8:159–168.

Miettinen OS, Cook EF (1981). Confounding: Essence and detection. *American Journal of Epidemiology* 1981; 114:593-603.

Neyman J (1923). On the Application of Probability Theory to Agricultural Experiments: Essay on Principles, Section 9. Translated in *Statistical Science* 1990; 5:465–480.

Ogburn EL, VanderWeele TJ (2012). On the nondifferential misclassification of a binary confounder. *Epidemiology*; 23(3):433-439.

Pearl J (1995). Causal diagrams for empirical research. *Biometrika*; 82:669–710.

Pearl J (2009). *Causality: Models, Reasoning, and Inference*, 2nd ed. New York: Cambridge University Press.

Pearl J, Robins JM (1995). Probabilistic evaluation of sequential plans from causal models with hidden variables. In: *Proceedings of the 11th Conference on Uncertainty in Artificial Intelligence*. Montreal, Quebec, Canada, pp. 444–453.

Pearson K, Lee A, Bramley-Moore L (1899). VI. Mathematical contributions to the Theory of Evolution.—VI. Genetic (Reproductive) selection: Inheritance of fertility in man, and of fecundity in thoroughbred horses. *Philosophical Transactions of the Royal Society of London, Series A* 192: 258–331.

Smith GCS, Pell JP (2003). Parachute use to prevent death and major trauma related to gravitational challenge: systematic review of randomised controlled trials. *British Medical Journal* 327:1459–1461.

Richardson TS, Robins JM (2010). Analysis of the binary instrumental variable model. In: Dechter R, Geffner H, Halpern JY, eds. *Heuristics, Probability and Causality: A Tribute to Judea Pearl*. College Publications, UK.

Richardson TS, Robins JM (2013). Single world intervention graphs (SWIGs): A unification of counterfactual and graphical approaches to causality. Working Paper Number 128. Available at http://www.csss.washington.edu/Papers/

Robins JM (1986). A new approach to causal Inference in mortality studies with sustained exposure periods –Application to control of the healthy worker survivor effect. *Mathematical Modelling* 7:1393–1512 (errata appeared in *Computers and Mathematics with Applications* 1987;14:917–921).

Robins JM (1987). Addendum to "A new approach to causal inference in mortality studies with sustained exposure periods –Application to control of the healthy worker survivor effect". *Computers and Mathematics with Applications* 14 (9–12):923–945 (errata appeared in *Computers and Mathematics with Applications* 1987;18:477).

Robins JM (1988). Confidence intervals for causal parameters. *Statistics in Medicine* 7:773–785.

Robins JM (1989). The analysis of randomized and non-randomized AIDS treatment trials using a new approach to causal inference in longitudinal studies. In: Sechrest L, Freeman H, Mulley A, eds. *Health Services Research Methodology: A Focus on AIDS*. NCHRS, U.S. Public Health Service; 113–159.

Robins JM. (1994). Correcting for non-compliance in randomized trials using structural nested mean models. *Communications in Statistics* 23:2379–2412.

Robins JM (1997). Causal Inference from Complex Longitudinal Data. *Latent Variable Modeling and Applications to Causality*. Berkane M, ed. New York, NY: Springer Verlag, pp. 69–117.

Robins JM (1998). Marginal structural models. *1997 Proceedings of the Section on Bayesian Statistical Science*. Alexandria, Virginia: American Statistical Association, 1–10.

Robins JM (2003). Semantics of causal DAG models and the identification of direct and indirect effects. In: Green P, Hjort NL, Richardson S, eds. *Highly Structured Stochastic Systems*. New York: Oxford University Press; p. 70–81.

Robins JM (2000). Robust estimation in sequentially ignorable missing data and causal inference models. *1999 Proceedings of the Section on Bayesian Statistical Science*. Alexandria, Virginia: American Statistical Association, pp. 6–10.

Robins JM, Greenland S (1989). Estimability and estimation of excess and etiologic fraction. *Statistics in Medicine* 8:845–859.

Robins JM, Greenland S (2000). Comment on "Causal inference without counterfactuals." *Journal of the American Statistical Association* 95:477–82.

Robins JM, Hernán MA, Rotnitzky A (2007). Effect modification by time-varying covariates. *American Journal of Epidemiology* 166:994-1002.

Robins JM, Morgenstern H (1987). The foundations of confounding in epidemiology. Computers & Mathematics with Applications 14(9-12); 869–916.

Robins JM, Richardson TS (2011). Alternative graphical causal models and the identification of direct effects. In: *Causality and Psychopathology: Finding the Determinants of Disorders and Their Cures*. P. Shrout, ed. New York, NY: Oxford University Press.

Robins JM, Rotnitzky A, Scharfstein D (1999). Sensitivity analysis for selection bias and unmeasured confounding in missing data and causal inference models. In: *Statistical Models in Epidemiology: The Environment and Clinical Trials*. Halloran ME, Berry D, eds. New York: Springer-Verlag, pp. 1–92.

Rosenbaum PR (1987). Model-based direct adjustment. *Journal of the American Statistical Association* 82:387–394.

Rosenbaum PR (2007). Interference between units in randomized experiments. *Journal of the American Statistical Association* 102:191–200.

Rosenbaum PR (2002). *Observational Studies*, 2nd edition. New York, NY: Springer-Verlag.

Rosenbaum PR, Rubin DB (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika* 70:41–55.

Rothman KJ (1976). Causes. *American Journal of Epidemiology* 104:587–592.

Rothman KJ, Greenland S (1998). *Modern Epidemiology*, 2nd edition. Philadelphia, PA: Lippincott Williams and Wilkins.

Rothman KJ, Greenland S, Walker AM (1980). Concepts of interaction. *American Journal of Epidemiology* 112:467–470.

Rothman KJ, Greenland S, Poole C, Lash TL (2008). Causation and Causal Inference. In: Rothman KJ, Greenland S, Lash TL, eds. *Modern Epidemiology*, 3rd ed. Philadelphia, PA: Lippincott Williams and Wilkins.

Rubin DB (1974). Estimating causal effects of treatments in randomized and nonrandomized studies. *Journal of Educational Psychology* 56:688–701.

Rubin DB (1976). Inference and missing data (with discussion). *Biometrika* 63:581–592.

Rubin DB (1978). Bayesian inference for causal effects: The role of randomization. *Annals of Statistics* 6:34–58.

Rubin DB (1980). Discussion of "Randomized analysis of experimental data: the Fisher randomization test" by Basu D. *Journal of the American Statistical Association* 75:591–593.

Samuels ML (1981). Matching and design efficiency in epidemiological studies. *Biometrika* 68:577–588.

Saracci R (1980). Interaction and synergism. *American Journal of Epidemiology* 112:465–466.

Sato T, Matsuyama Y (2003). Marginal structural models as a tool for standardization. *Epidemiology* 14:680–686.

Simpson EH (1951). The interpretation of interaction in contingency tables. *Journal of the Royal Statistical Society, Series B* 13:238–241.

Spirtes P, Glymour C, Scheines R (2000). *Causation, Prediction and Search*, 2nd ed. Cambridge, MA: MIT Press.

Sobel ME (2006). What do randomized studies of housing mobility demonstrate? *Journal of the American Statistical Association* 101: 1398–1407.

Stalnaker RC (1968). A theory of conditionals. In Rescher N, ed. *Studies in Logical Theory*. Oxford: Blackwell. Reprinted in Jackson F, ed. *Conditionals*. Oxford: Oxford University Press, 1991.

Stuart EA (2010). Matching methods for causal inference. *Statistical Science* 25, 1–21.

Tchetgen Tchetgen EJ, Rotnitzky A (2011). Double-robust estimation of an exposure-outcome odds ratio adjusting for confounding in cohort and case-control studies. *Statistics in Medicine* 30(4):335–47.

Toh SW, Hernán MA (2008). Causal inference from longitudinal studies with baseline randomization. *International Journal of Biostatistics* ; 4: Article 22. Available at http://www.bepress.com/ijb/vol4/iss1/22.

VanderWeele TJ, Hernán MA (2006). From counterfactuals to sufficient component causes and vice versa. *European Journal of Epidemiology* 21:855–858.

VanderWeele TJ (2009). On the distinction between interaction and effect modification. *Epidemiology* 20:863–871.

VanderWeele TJ (2009). Concerning the consistency assumption in causal inference. *Epidemiology* 20:880–883.

VanderWeele TJ (2010). Empirical tests for compositional epistasis. *Nature Reviews Genetics* 11:166.

VanderWeele TJ, Hernán MA, Robins JM (2008). Causal directed acyclic graphs and the direction of unmeasured confounding bias. *Epidemiology* 19:720–728.

VanderWeele TJ, Robins JM (2007a). The identification of synergism in the sufficient-component-cause framework. *Epidemiology* 18:329–339.

VanderWeele TJ, Robins JM (2007b). Four types of effect modification. A classification based on directed acyclic graphs. *Epidemiology* 18; 561–568.

VanderWeele TJ, Robins JM (2007c). Directed acyclic graphs, sufficient causes, and the properties of conditioning on a common effect. *American Journal of Epidemiology* 166; 1096–1104.

VanderWeele TJ, Robins JM (2008). Empirical and counterfactual conditions for sufficient cause interactions. *Biometrika* 95:49–61.

VanderWeele TJ, Hernán MA (2009). Results on differential and dependent measurement error of the exposure and

the outcome using signed directed acyclic graphs. *American Journal of Epidemiology* 175(12):1303–1310.

VanderWeele TJ, Shpitser I (2013). On the definition of a confounder. *Annals of Statistics* 41(1): 196–220.

Walker AM (1991). *Observation and Inference: An Introduction to the Methods of Epidemiology.* Newton Lower Falls, MA: Epidemiology Resources, Inc.

Wasserman L (2004). *All of Statistics: A Concise Course in Statistical Inference.* New York, NY: Springer.

Weinberg CR, Umbach DM, Greenland S (1994). When will nondifferential misclassification of an exposure preserve the direction of a trend? *American Journal of Epidemiology* 140:565–571.

Yule GU (1903). Notes on the theory of association of attributes in statistics. *Biometrika* 2:121–134.